

## SEQUENCE LISTING

5

## GENERAL INFORMATION

10 <110> APPLICANT: Lingappa, Jaisri  
Lingappa, Vishwanath

15 <120> TITLE OF THE INVENTION: HIV Capsid Assembly Associated  
Compositions and Method

<130> FILE REFERENCE LING.001.01US

<140> CURRENT APPLICATION NUMBER:

15 <141> CURRENT FILING DATE: 2002-01-02

## CORRESPONDENCE ADDRESS:

20 ADDRESSEE: Rae-Venter Law Group, P.C.  
STREET: 260 Sheridan Avenue, Suite 440, PO Box 60039  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94306

## ATTORNEY/AGENT INFORMATION:

25 NAME: Rae-Venter, Barbara  
REGISTRATION NUMBER: 32,750

## TELECOMMUNICATION INFORMATION:

30 TELEPHONE: 650-328-4400  
TELEFAX: 650-328-4477

## COMPUTER READABLE FORM:

35 MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: PatentIn version 3.1

## PRIOR APPLICATION DATA:

40 APPLICATION NUMBER: 09/020,144  
FILING DATE: 06-FEB-1998

<160>NUMBER OF SEQUENCES: 6

<170> PatentIn Version 3.0

<210>INFORMATION FOR SEQ ID NO:1

5

SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1610 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

10

- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(C) ISOLATE: DNA coding sequence for HIV capsid protein Pr55

15

<400> SEQUENCE DESCRIPTION: SEQ ID NO:1

ATGGGTGCGA	GAGCGTCGGT	ATTAAGCGGG	GGAGAATTAG	ATAAATGGGA	AAAAATTCTGG	60
TTAAGGCCAG	GGGGAAAGAA	AAAATATAAG	TTAAAACATA	TAGTATGGC	AAGCAGGGAG	120
CTAGAACGAT	TCGCAGTC	AA TCCTGGCCTG	TTAGAACAT	CAGAAGGCTG	CAGACAAATA	180
TTGGGACAGC	TACAGCCATC	CCTTCAGACA	GGATCAGAAG	AACTTAGATC	ATTATATAAT	240
ACAGTAGCAA	CCCTCTATTG	TGTACATCAA	AGGATAGATG	TAAAAGACAC	CAAGGAAGCT	300
TTAGAGAAGA	TAGAGGAAGA	GCAAAACAAA	AGTAAGAAAA	AGGCACAGCA	AGCAGCAGCT	360
GCAGCTGGCA	CAGGAAACAG	CAGCCAGGTC	AGCCAAAATT	ACCTATAGT	GCAGAACCTA	420
CAGGGGCAAA	TGGTACATCA	GGCCATATCA	CCTAGAACTT	TAAATGCATG	GGTAAAAGTA	480
GTAGAGAAGA	AGGCTTCAG	CCCAGAAGTA	ATACCCATGT	TTTCAGCATT	ATCAGAAGGA	540
GCCACCCAC	AAGATTAAA	CACCATGCTA	AACACAGTGG	GGGGACATCA	AGCAGCCATG	600
CAAATGTTAA	AAGAGACTAT	CAATGAGGAA	GCTGCAGAAT	GGGATAGAGT	GCATCCAGTG	660
CATGCAGGGC	CTATTGCACC	AGGCCAAATG	AGAGAACAA	GGGAACTGA	CATAGCAGGA	720
30	ACTACTAGTA	CCCTTCAGGA	ACAAATAGGA	TGGATGACAA	ATAATCCACC	780
	GGAGAAATCT	ATAAAAGATG	GATAATCCTG	GGATTAAATA	AAATAGTAAG	840
	CCTACCAGCA	TTCTGGACAT	AAGACAAGGA	CCAAAGGAAC	CCTTTAGAGA	900
	CGGTTCTATA	AAACTCTAAG	AGCCGAACAA	GCTTCACAGG	ATGTAAAAAA	960
	GAAACCTTGT	TGGTCCAAAA	TGCAAACCCA	GATTGTAAGA	CTATTAAAGA	1020
35	CCAGCAGCTA	CACTAGAAGA	AATGATGACA	GCATGTCAGG	GAGTGGGGGG	1080
	AAAGCAAGAG	TTTTGGCTGA	AGCCATGAGC	CAAGTAACAA	ATCCAGCTAA	1140
	CAGAGAGGCA	ATTTTAGGAA	CCAAAGAAAG	ACTGTTAAGT	GTTTCAATTG	1200
	GGGCACATAG	CCAAAAATTG	CAGGGCCCT	AGGAAAAGG	GCTGTTGGAG	1260
	GAAGGACACC	AAATGAAAGA	TTGCACTGAG	AGACAGGCTA	ATTTTTAGG	1320
40	CCTTCCTACA	AGGGAAGGCC	AGGGAATTTC	CTTCAGAGCA	GACCAGAGCC	1380
	CCAGAAGAGA	GCTTCAGGTT	TGGGGAGGAG	AAAACAACTC	CCTCTCAGAA	1440

ATAGACAAGG AACTGTATCC TTTAACTTCC CTCAGATCAC TCTTTGGCAA CGACCCCTCG 1500  
TCACAATAAG GATAGGGGGG CAACTAAAGG AAGCTCTATT AGATACAGGA GCAGATGATA 1560  
CAGTATTAGA AGAAATGAAT TTGCCAGGAA AATGGAAACC AAAAATGATA 1610

5 <210> INFORMATION FOR SEQ ID NO:2

SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 amino acids  
(B) TYPE: amino acid  
10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: PRT

(vi) ORIGINAL SOURCE:

(C) ISOLATE: peptide fragment of host cell (wheat germ) protein

15 HP68

<400> SEQUENCE DESCRIPTION: SEQ ID NO:2

Pro Arg Pro Tyr Leu Asp Val Lys Gln Arg Leu Lys Ala Ala Arg Val  
1 5 10 15  
Ile Arg Ser Leu Leu Arg Ser Asn  
20

<210> INFORMATION FOR SEQ ID NO:3

SEQUENCE CHARACTERISTICS:

(A) LENGTH: 44 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(C) ISOLATE: Degenerate oligonucleotide C-terminal peptide

35 sequence of WGHP68

<400> SEQUENCE DESCRIPTION: SEQ ID NO:3

ATGAATTAC TGGGACTGCG GATAGATTAC TGGTACTGGG GATC 44

40

<210> INFORMATION FOR SEQ ID NO:4

## SEQUENCE CHARACTERISTICS:

5 (A) LENGTH: 44 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

10 (C) ISOLATE: Degenerate oligonucleotide C-terminal peptide sequence of WGHP68

<400> SEQUENCE DESCRIPTION: SEQ ID NO:4

15 ATGAATTCAC TGGGCTCTGA TAGATTACTG GTACTGGGA TC 42

<210> SEQ ID NO:5

<211> Length:604

<212> Type: PRT

<213> Organism: Triricum aestivum

<400> Sequence 5

Met Ala Asp Arg Leu Thr Arg Ile Ala Ile Val Ser Glu Asp Lys Cys

1 5 10 15

Lys Pro Lys Lys Cys Arg Gln Glu Cys Lys Lys Ser Cys Pro Val Val

20 25 30

Lys Thr Gly Lys Leu Cys Ile Glu Val Ser Pro Val Ala Lys Leu Ala  
35 40 45

30 Phe Ile Ser Glu Glu Leu Cys Ile Gly Cys Gly Ile Cys Val Lys Lys  
50 55 60

Cys Pro Phe Asp Ala Ile Glu Ile Asn Leu Pro Lys Asp Leu Glu  
65 70 75 80

Lys Asp Thr Thr His Arg Tyr Gly Pro Asn Thr Phe Lys Leu His Arg  
85 90 95

Leu Pro Val Pro Arg Pro Gly Gln Val Leu Gly Leu Val Gly Thr Asn  
100 105 110

Gly Ile Gly Lys Ser Thr Ala Leu Lys Val Leu Ala Gly Lys Leu Lys  
115 120 125

40 Pro Asn Leu Gly Arg Phe Lys Asn Pro Pro Asp Trp Gln Glu Ile Leu  
130 135 140

Thr Tyr Phe Arg Gly Ser Glu Leu Gln Asn Tyr Phe Thr Arg Ile Leu  
145 150 155 160  
Glu Asp Asn Leu Lys Ala Ile Ile Lys Pro Gln Tyr Val Asp His Ile  
165 170 175  
5 Pro Lys Ala Val Gln Gly Asn Val Gly Gln Val Leu Glu Gln Lys Asp  
180 185 190  
Glu Arg Asp Met Lys Asn Glu Leu Cys Val Asp Leu Glu Leu Asn Gln  
195 200 205  
Val Ile Asp Arg Asn Val Gly Asp Leu Ser Gly Gly Glu Leu Gln Arg  
10 210 215 220  
Phe Ala Ile Ala Val Val Ala Val Gln Ser Ala Glu Ile Tyr Met Phe  
225 230 235 240  
Asp Glu Pro Ser Ser Tyr Leu Asp Val Lys Gln Arg Leu Lys Ala Ala  
245 250 255  
15 Arg Val Ile Arg Ser Leu Leu Arg Ser Asn Ser Tyr Val Ile Val Val  
260 265 270  
Glu His Asp Leu Ser Val Leu Asp Tyr Leu Ser Asp Phe Ile Cys Cys  
275 280 285  
Leu Tyr Gly Lys Pro Gly Ala Tyr Gly Val Val Thr Leu Pro Phe Ser  
20 290 295 300  
Val Arg Glu Gly Ile Asn Ile Phe Leu Ala Gly Phe Val Pro Thr Glu  
305 310 315 320  
Asn Leu Arg Phe Arg Asp Glu Ser Leu Thr Phe Lys Ile Ala Glu Thr  
325 330 335  
25 Gln Glu Ser Ala Glu Glu Val Ala Thr Tyr Gln Arg Tyr Lys Tyr Pro  
340 345 350  
Thr Met Ser Lys Thr Gln Gly Asn Phe Lys Leu Ser Val Val Glu Gly  
355 360 365  
Glu Phe Thr Asp Ser Gln Ile Val Val Met Leu Gly Glu Asn Gly Thr  
30 370 375 380  
Gly Lys Thr Thr Phe Ile Arg Met Leu Ala Gly Leu Leu Lys Pro Asp  
385 390 395 400  
Thr Met Glu Gly Thr Glu Val Glu Ile Pro Glu Phe Asn Val Ser Tyr  
405 410 415  
35 Lys Pro Gln Lys Ile Ser Pro Lys Phe Gln His Pro Val Arg His Leu  
420 425 430  
Leu His Ser Lys Ile Arg Asp Ser Tyr Thr His Pro Gln Phe Val Ser  
435 440 445  
Asp Val Met Lys Pro Leu Gln Ile Glu Gln Leu Met Asp Gln Glu Val  
40 450 455 460  
Ile Asn Leu Ser Gly Gly Glu Leu Gln Arg Val Ala Leu Cys Leu Cys

465 470 475 480  
Leu Gly Lys Pro Ala Asp Ile Tyr Leu Ile Asp Glu Pro Ser Ala Tyr  
485 490 495  
Leu Asp Ser Glu Gln Arg Ile Val Ala Ser Lys Val Ile Lys Arg Phe  
5 500 505 510  
Ile Leu His Ala Lys Lys Thr Ala Phe Ile Val Glu His Asp Phe Ile  
515 520 525  
Met Ala Thr Tyr Leu Ala Asp Lys Val Ile Val Tyr Glu Gly Leu Ala  
530 535 540  
10 Ser Ile Asp Cys Thr Ala Asn Ala Pro Gln Ser Leu Val Ser Gly Met  
545 550 555 560  
Asn Lys Phe Leu Ser His Leu Asp Ile Thr Phe Arg Arg Asp Pro Thr  
565 570 575  
Asn Tyr Arg Pro Arg Ile Asn Lys Leu Glu Ser Thr Lys Asp Arg Glu  
15 580 585 590  
Gln Lys Asn Ala Gly Ser Tyr Tyr Tyr Leu Asp Asp  
595 600

<210> SEQ ID NO:6

<211> LENGTH: 599

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE 6

25 Met Ala Asp Lys Leu Thr Arg Ile Ala Ile Val Asn His Asp Lys Cys  
1 5 10 15  
Lys Pro Lys Lys Cys Arg Gln Glu Cys Lys Lys Ser Cys Pro Val Val  
20 25 30  
Arg Met Gly Lys Leu Cys Ile Glu Val Thr Pro Gln Ser Lys Ile Ala  
30 35 40 45  
Trp Ile Ser Glu Thr Leu Cys Ile Gly Cys Gly Ile Cys Ile Lys Lys  
50 55 60  
Cys Pro Phe Gly Ala Leu Ser Ile Val Asn Leu Pro Ser Asn Leu Glu  
65 70 75 80  
35 Lys Glu Thr Thr His Arg Tyr Cys Ala Asn Ala Phe Lys Leu His Arg  
85 90 95  
Leu Pro Ile Pro Arg Pro Gly Glu Val Leu Gly Leu Val Gly Thr Asn  
100 105 110  
Gly Ile Gly Lys Ser Ala Ala Leu Lys Ile Leu Ala Gly Lys Gln Lys  
40 115 120 125  
Pro Asn Leu Gly Lys Tyr Asp Asp Pro Pro Asp Trp Gln Glu Ile Leu

130 135 140  
Thr Tyr Phe Arg Gly Ser Glu Leu Gln Asn Tyr Phe Thr Lys Ile Leu  
145 150 155 160  
Glu Asp Asp Leu Lys Ala Ile Ile Lys Pro Gln Tyr Val Ala Arg Phe  
5 165 170 175  
Leu Arg Leu Ala Lys Gly Thr Val Gly Ser Ile Leu Asp Arg Lys Asp  
180 185 190  
Glu Thr Lys Thr Gln Ala Ile Val Cys Gln Gln Leu Asp Leu Thr His  
195 200 205  
10 Leu Lys Glu Arg Asn Val Glu Asp Leu Ser Gly Gly Glu Leu Gln Arg  
210 215 220  
Phe Ala Cys Ala Val Val Cys Ile Gln Lys Ala Asp Ile Phe Met Phe  
225 230 235 240  
Asp Glu Pro Ser Ser Tyr Leu Asp Val Lys Gln Arg Leu Lys Ala Ala  
15 245 250 255  
Ile Thr Ile Arg Ser Leu Ile Asn Pro Asp Arg Tyr Ile Ile Val Val  
260 265 270  
Glu His Asp Leu Ser Val Leu Asp Tyr Leu Ser Asp Phe Ile Cys Cys  
275 280 285  
20 Leu Tyr Gly Val Pro Ser Ala Tyr Gly Val Val Thr Met Pro Phe Ser  
290 295 300  
Val Arg Glu Gly Ile Asn Ile Phe Leu Asp Gly Tyr Val Pro Thr Glu  
305 310 315 320  
Asn Leu Arg Phe Arg Asp Ala Ser Leu Val Phe Lys Val Ala Glu Thr  
325 330 335  
25 Ala Asn Glu Glu Val Lys Lys Met Cys Met Tyr Lys Tyr Pro Gly  
340 345 350  
Met Lys Lys Lys Met Gly Glu Phe Glu Leu Ala Ile Val Ala Gly Glu  
355 360 365  
30 Phe Thr Asp Ser Glu Ile Met Val Met Leu Gly Glu Asn Gly Thr Gly  
370 375 380  
Lys Thr Thr Phe Ile Arg Met Leu Ala Gly Arg Leu Lys Pro Asp Glu  
385 390 395 400  
Gly Gly Glu Val Pro Val Leu Asn Val Ser Tyr Lys Pro Gln Lys Ile  
35 405 410 415  
Ser Pro Lys Ser Thr Gly Ser Val Arg Gln Leu Leu His Glu Lys Ile  
420 425 430  
Arg Asp Ala Tyr Thr His Pro Gln Phe Val Thr Asp Val Met Lys Pro  
435 440 445  
40 Leu Gln Ile Glu Asn Ile Ile Asp Gln Glu Val Gln Thr Leu Ser Gly  
450 455 460

Gly Glu Leu Gln Arg Val Arg Leu Arg Leu Cys Leu Gly Lys Pro Ala  
465 470 475 480  
Asp Val Tyr Leu Ile Asp Glu Pro Ser Ala Tyr Leu Asp Ser Glu Gln  
485 490 495  
5 Arg Leu Met Ala Ala Arg Val Val Lys Arg Phe Ile Leu His Ala Lys  
500 505 510  
Lys Thr Ala Phe Val Val Glu His Asp Phe Ile Met Ala Thr Tyr Leu  
515 520 525  
Ala Asp Arg Val Ile Val Phe Asp Gly Val Pro Ser Lys Asn Thr Val  
10 530 535 540  
Ala Asn Ser Pro Gln Thr Leu Leu Ala Gly Met Asn Lys Phe Leu Ser  
545 550 555 560  
Gln Leu Glu Ile Thr Phe Arg Arg Asp Pro Asn Asn Tyr Arg Pro Arg  
565 570 575  
15 Ile Asn Lys Leu Asn Ser Ile Lys Asp Val Glu Gln Lys Lys Ser Gly  
580 585 590  
Asn Tyr Phe Phe Leu Asp Asp  
595

20  
25

30

35

WHAT IS CLAIMED IS: